

BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1 Tinjauan Pustaka

Dalam penelitian ini digunakan beberapa sumber pustaka. Pustaka yang relevan pada penelitian ini ditinjau dari sisi kasus penelitian dan metode yang digunakan. Kasus penelitian yang dilakukan adalah mengenai klasifikasi status gizi mahasiswa. Metode yang digunakan dalam penelitian ini adalah metode *Naïve Bayes Classification*.

Beberapa penelitian yang dilakukan terkait dengan klasifikasi gizi yang menggunakan metode *K-Nearest Neighbor* yaitu Herwanto (2015), Hasani, dkk, (2015), dapat diterangkan bahwa perhitungan klasifikasi gizi menggunakan *K-Nearest Neighbor* sama dengan perhitungan manual menggunakan *excel*.

Penelitian selanjutnya yang dilakukan terkait dengan klasifikasi gizi yang menggunakan metode *Naïve bayes classification* yaitu Ahmad Saleh (2015), Muhamad Natsir (2016), Sri Kusumadewi (2015), dapat diterangkan bahwa kasifikasi gizi yang menggunakan metode *Naïve bayes classification* hasilnya menunjukan metode *Naïve Bayes Classificatin* dapat memecahkan masalah dengan cukup baik dengan kinerja sistem sebesar 93%. Dimana perhitungan probabilistik kemunculan hasil diagnosa gizi balita berdasarkan pada perhitungan algoritma *naïve bayes* dalam menentukan apakah balita memiliki gizi lebih, baik, kurang dan buruk.

Pada penelitian yang akan penulis lakukan, diharapkan dapat mencapai hasil perhitungan klasifikasi Algoritma *Naïve bayes classification* status gizi berdasarkan jenis kelamin sesuai dengan standarisasi dari kementrian kesehatan.

Tabel 2.1 Perbandingan Hasil Penelitian

Nama Peneliti	Objek	Metode Data Mining	Hasil
Herwanto, (2015)	Anak Balita	<i>K-Nearest Neighbor</i>	Perbandingan tertinggi yaitu pada saat pembobotan bernilai k sama dengan 3 dengan tingkat bernilai k sama dengan sebesar 93%
Ahmad Saleh, (2015)	Manusia	<i>Naïve Bayes</i>	Aplikasi pola penggunaan metode klasifikasi <i>naïve bayes</i> pada penelitian ini yaitu 94,12 % dan kesalahan sebesar 5,88 % dari 119 data
Muhamad Natsir (2016)	Anak Balita	<i>Naïve Bayes</i>	Diagnosa gizi balita dan pemantauan kondisi dari balita hingga berumur 5 tahun
Sri Kusumadewi (2009)	Remaja putri	<i>Naïve Bayes Classification</i>	Klasifikasi status gizi berdasarkan hasil pengukuran antropometri dan model sistem yang dibangun memiliki kinerja yang baik karena hasil pengujian menunjukkan total kinerja sebesar 93,2%.
Hasani, dkk, (2015)	Mahasiswa	<i>K-Nearest Neighbor</i>	Klasifikasi status gizi mahasiswa
Dhimas Tantra Putra (2015)	Orang dewasa (puskesmas jiken)	<i>Naïve Bayes Classification</i>	Berdasarkan pengujian akurasi yang dilakukan, sistem status gizi orang dewasa dengan metode <i>naïve bayes classification</i> mempunyai tingkat akurasi sebesar 87,91 %.
Yang diusulkan Nurhayati Ibrahim (2017)	Mahasiswa	<i>Naïve Bayes Classification</i>	Diharapkan dengan penelitian ini dapat memperoleh hasil dengan akurasi sistem sebesar 90%.

Dari hasil penelitian sebelumnya dapat disimpulkan bahwa hasil atau akurasi dari setiap sistem yang dibuat oleh para peneliti sebelumnya memperoleh hasil akurasi sistem diatas 90%. Jadi kesimpulannya yaitu sistem yang dibuat memperoleh hasil yang baik.

2.2 Dasar Teori

2.2.1 Status Gizi

Status gizi adalah status gizi kesehatan yang dihasilkan oleh keseimbangan antara kebutuhan dan masukan nutrisi (Back 2002 dalam Jafar 2010). Menurut Almatsier (2005) status gizi didefinisikan sebagai suatu keadaan tubuh akibat konsumsi makanan dan penggunaan zat-zat gizi. Kekurangan gizi atau kelebihan gizi biasanya disebut *malnutrisi*. *Malnutrisi* merupakan keadaan patologis akibat kekurangan atau kelebihan zat gizi, baik secara relatif maupun absolut. Malnutrisi dapat terjadi karena kekurangan gizi (*undernutrition*) maupun kelebihan gizi (*overnutrition*). Keduanya disebabkan oleh ketidak seimbangan antara kebutuhan tubuh dan asupan zat gizi *esensial*. *Malnutrisi* adalah istilah umum untuk suatu kondisi medis yang disebabkan oleh pemberian atau cara makan yang tidak tepat atau tidak mencukupi. Pada dasarnya, konsumsi makanan bertujuan untuk mencapai status gizi optimal. Upaya penyediaan pangan agar tercapai status gizi optimal dapat dilakukan dengan mengkonsumsi karbohidrat, lemak, protein, vitamin, dan mineral sesuai dengan angka kecukupan gizi dalam rangka proses energi, transformasi, dan interaksinya dengan zat lain demi tercapainya keseimbangan energi tubuh.

2.2.2 Pengertian IMT

Antropometri merupakan cara penentuan status gizi yang paling mudah dan murah. Pengukuran *antropometri* adalah pengukuran yang digunakan untuk menentukan keadaan gizi seseorang, agar memperoleh hasil yang tepat, diberikan suatu patokan sebagai pedoman. Pengukuran *antropometri* untuk usia dewasa sekarang ini menggunakan perhitungan IMT. Indeks Massa Tubuh (IMT) adalah perbandingan (rasio) berat badan / tinggi badan yang sering digunakan untuk menilai status gizi orang dewasa. Dengan rumus IMT sebagai berikut:

$$IMT = \frac{\text{Berat Badan (kg)}}{[\text{Tinggi Badan (m)}]^2} \dots\dots\dots(2.1)$$

Beberapa studi telah mengungkapkan bahwa IMT adalah alat pengukuran yang berguna untuk mengukur obesitas, dan telah direkomendasikan untuk evaluasi klinik (et al, 1997).

IMT tidak mengukur lemak tubuh secara langsung, tetapi penelitian menunjukkan bahwa IMT berkorelasi dengan pengukuran secara langsung lemak tubuh seperti *underwater weighing* dan *dual energy x-ray absorbtometry* (Grummer-Strawn LM et al., 2002).

A. Kategori Indeks Massa Tubuh (IMT)

Untuk orang dewasa yang berusia 20 tahun keatas, IMT diinterpretasi menggunakan kategori status berat badan standar yang sama untuk semua umur bagi pria dan wanita. Untuk anak-anak, interpretasi IMT adalah spesifik mengikuti usia dan jenis kelamin (CDC, 2009).

Secara umum, IMT 25 keatas membawa arti pada obesitas. Standar baru untuk IMT telah dipublikasikan pada tahun 1998 mengklasifikasikan IMT di

bawah 18,5 sebagai sangat kurus atau *underweight*, IMT melebihi 23 sebagai berat badan lebih atau *overweight*, dan IMT melebihi 25 sebagai obesitas. IMT yang ideal bagi orang dewasa adalah diantara 18,5 hingga 22,9. Obesitas dikategorikan pada tiga tingkat I (25-29,9), tingkat II (30-40), dan tingkat III (>40) (WHO, 2004).

Untuk kepentingan Indonesia, batas ambang dimodifikasi lagi berdasarkan pengalaman klinis dan hasil penelitian di beberapa Negara. Batas ambang indeks massa tubuh (IMT) untuk Indonesia adalah sebagai berikut:

Tabel 2.2 Batas Ambang IMT Indonesia (Depkes, 2003)

Gender	Kategori IMT (Kg/m ²)			
	Kurus	Normal	Obesitas Ringan	Obesitas Berat
Pria	IMT < 18 kg/m ²	18 ≤ IMT < 25 kg/m ²	≥ 25 IMT < 27 kg/m ²	≥ 27 IMT kg/m ²
Wanita	IMT < 17 kg/m ²	17 ≤ IMT < 23 kg/m ²	≥ 23 IMT < 27 kg/m ²	≥ 27 IMT kg/m ²

Keterangan :

IMT < 17 : keadaan orang tersebut disebut kurus

IMT > 18-25 : keadaan orang tersebut disebut normal

IMT > 23 - < 27 : keadaan orang tersebut disebut obesitas tingkat ringan

IMT > 27 : keadaan orang tersebut disebut obesitas tingkat berat

a. Indeks Massa Tubuh (IMT) Kurus

Indeks massa tubuh dikategorikan kurus jika pembagian berat per kuadrat tingginya kurang dari 18 kg/m^2 . Penyebab rata-rata dikarenakan konsumsi energi lebih rendah dari kebutuhan yang mengakibatkan sebagian cadangan energi tubuh dalam bentuk lemak akan digunakan. Kerugiannya jika seseorang masuk dalam kategori ini antara lain :

1. Penampilan cenderung kurang menarik
2. Mudah letih
3. Resiko sakit tinggi, beberapa resiko sakit yang dihadapi antara lain :
penyakit infeksi, depresi, anemia dan diare

b. Indeks Massa Tubuh (IMT) Normal

Indeks massa tubuh termasuk normal jika pembagian berat per kuadrat tertingginya antara 18 sampai 25 kg/m^2 . Kategori ini bisa diwujudkan dengan mengkonsumsi energi sesuai dengan jumlah yang dibutuhkan tubuh. Sehingga terjadi penimbunan energi dalam bentuk lemak, maupun penggunaan lemak sebagai sumber energi. Keuntungan dari IMT yang normal ini antara lain :

1. Penampilan menarik, proposional, dan lincah
2. Resiko penyakit bisa di minimalisir menjadi lebih rendah

c. Indeks Massa Tubuh (IMT) Obesitas

Obesitas berpotensi menjadi faktor primer kasus *degenerative* dan *metabolic sindrom*. Beberapa studi menunjukkan bahwa obesitas adalah resiko yang paling tinggi untuk penyakit jantung, DM, dan beberapa jenis kanker. Adapun kerugian atau resiko dari kategori ini adalah :

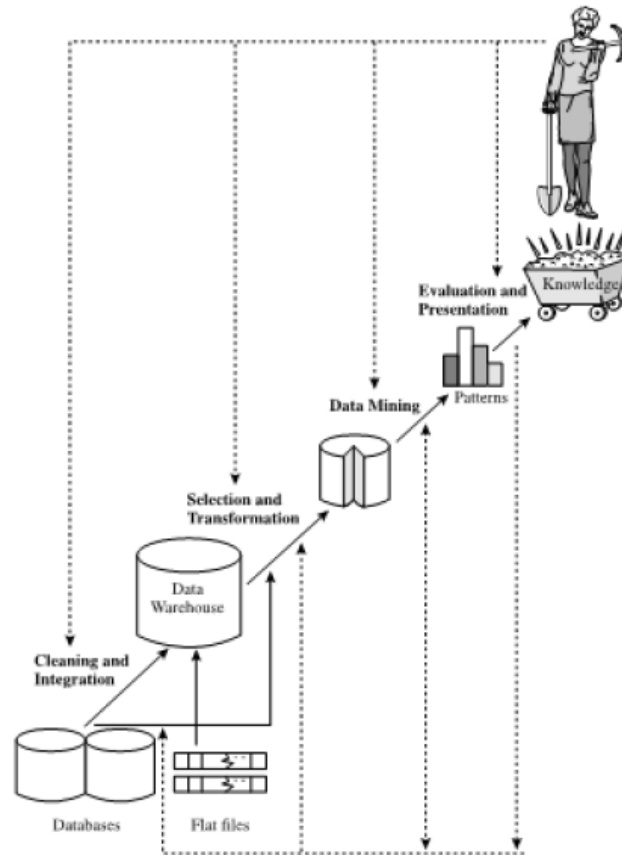
1. Penampilan kurang menarik
2. Gerakan tidak gesit dan terlambat
3. Merupakan faktor resiko penyakit : jantung dan pembuluh darah, kencing manis (*diabetes melitus*), tekanan darah tinggi, gangguan sendi dan tulang (*degeneratif*), gangguan fungsi ginjal, kanker, pada wanita dapat mengakibatkan gangguan haid (haid tidak teratur) (Chrlotte, 2000).

2.2.3 Klasifikasi

Klasifikasi merupakan salah satu tugas yang penting dalam data *mining*, mengorganisasikan serta mengelompokkan data kedalam kelas-kelas yang berbeda merupakan tujuan utama dari klasifikasi. Definisi pengklasifikasian adalah sebuah fungsi yang bersifat prediksi dan menggolongkan data item tertentu ke dalam sebuah kelas. Sebuah pengklasifikasian dibuat dari sekumpulan data *training*/latih dengan kelas yang telah ditentukan dan dikenal ciri-cirinya sebelumnya. Kinerja pengklasifikasian biasanya diukur dengan ketepatan (Adilla, 2016).

2.2.4 Tahap-Tahap Data Mining

Sebagai suatu rangkaian proses, data *mining* dapat dibagi menjadi beberapa tahap proses yang diilustrasikan pada Gambar 1.1 tahap-tahap tersebut bersifat interaktif, pemakai terlibat langsung atau dengan perantaraan *knowledgebase*.



Gambar 2.1 Tahap-Tahap Data Mining

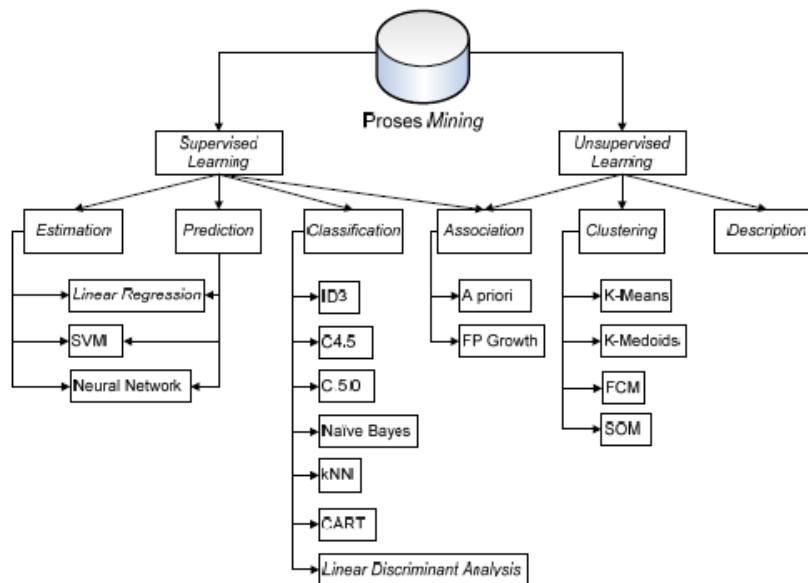
Tahap-tahap data *mining* adalah sebagai berikut:

- Pembersihan data (*data cleaning*) pembersihan data merupakan proses menghilangkan-kan *noise* dan data yang tidak konsisten atau data tidak relevan.
- Integrasi data (*data integration*)
Integrasi data merupakan penggabungan data dari berbagai *database* ke dalam satu *database* baru.
- Seleksi data (*data selection*)
Data yang ada pada database sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari *database*.
- Transformasi data (*data transformation*)

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data *mining*.

- Proses *mining*

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data. Beberapa metode yang dapat digunakan berdasarkan pengelompokan data *mining* dapat dilihat pada gambar 2.2.



Gambar 2.2 Metode Data Mining

- Evaluasi pola (*pattern evaluation*)

Untuk mengidentifikasi pola-pola menarik kedalam *knowledge based* yang ditemukan.

- Presentasi pengetahuan (*knowledgepresentation*)

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna.

2.2.5 Algoritma *Naïve Bayes*

Naïve Bayes merupakan salah satu metode *machine learning* yang menggunakan perhitungan probabilitas. Algoritma ini memanfaatkan teori

probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya. *Naïve bayes clasification* adalah metode yang berdasarkan probabilitas dan *Theorema Bayesian* dengan asumsi bahwa setiap variabel bersifat bebas (*independence*) dan mengasumsikan bahwa keberadaan sebuah fitur (variabel) tidak ada kaitannya dengan keberadaan fitur (variabel) yang lain. *Naïve Bayes* adalah model penyederhanaan dari metode *bayes*. *Naïve Bayes* inilah yang digunakan didalam *mechine learning* sebagai metode untuk mendapatkan hipotesis untuk keputusan.

Salah satu penerapan *Teorema Bayes* dalam klasifikasi *Naïve Bayes* adalah:

$$P(X|Y) = \frac{P(Y)\pi_{i=1}^d P(X_i|Y)}{P(X)} \dots\dots\dots (2.2)$$

Dimana :

- $P(X|Y)$ adalah probabilitas data dengan vektor X pada kelas Y
- $P(Y)$ adalah probabilitas awal
- $\pi_{i=1}^d P(X_i|Y)$ adalah probabilitas independen kelas Y dari semua fitur dalam vektor X.

Umumnya, *Bayesian* mudah dihitung untuk fitur bertipe kategoris seperti pada kasus fitur “jenis kelamin” dengan nilai “Laki-laki dan Perempuan” namun untuk fitur numerik ada perlakuan khusus sebelum dimasukkan dalam *Naïve Bayes*. Caranya adalah :

- a. Melakukan diskreditasi pada setiap fitur kontinu dan mengganti nilai fitur kontinu tersebut dengan nilai interval diskret. Pendekatan ini dilakukan dengan menstransformasikan fitur kontinu ke dalam fitur ordinal.

- b. Mengasumsikan bentuk tertentu dari distribusi probabilitas untuk fitur kontinu dan memperkirakan parameter distribusi dengan data penelitian. Distribusi *gaussian* sering dipilih untuk merepresentasikan peluang kelas bersyarat untuk atribut kontinu. Distribusi dikarakterisasi dengan dua parameter yaitu mean μ , dan varian σ^2 . Untuk tiap kelas y_j , peluang kelas bersyarat untuk X_i adalah:

$$P(x_i = x_i | y_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}^2}} \exp \left(-\frac{(x_i - \mu_{ij})^2}{2\sigma_{ij}^2} \right) \dots\dots\dots (2.3)$$

Dimana :

- Parameter μ_{ij} dapat diestimasi berdasarkan sampel mean X_i (x) untuk seluruh *training*/latih rekor yang dimiliki kelas y_j .
- σ_{ij}^2 dapat diestimasi dari sampel varian (s^2) *training*/latih rekor tersebut.

Adapun persamaan yang digunakan untuk menghitung nilai rata – rata (*mean*) dapat dilihat sebagai berikut :

$$\mu = \sum_{i=1}^n x_i$$

Atau

$$\mu = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \dots\dots\dots (2.4)$$

dimana :

- μ : rata – rata hitung (*mean*)
- x_i : nilai sample ke - i
- n : jumlah sampel

Persamaan untuk menghitung nilai simpangan baku (standar deviasi) dapat dilihat sebagai berikut:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n - 1}} \dots\dots\dots(2.5)$$

di mana :

- σ : standar deviasi
- x_i : nilai x ke $-i$
- μ : rata-rata hitung
- n : jumlah sampel

Persamaan untuk menghitung nilai normalisasi dapat menggunakan WP (weighted product) untuk menghitung nilai akhir normalisasi dapat dilihat sebagai berikut:

$$v = \frac{(D1)}{((D1) + (D2) + (D3) + (D4))} \dots\dots\dots(2.6)$$

2.2.6 Web

World Wide Web atau *www* atau juga dikenal dengan *WEB* adalah salah satu layanan yang didapat oleh pemakai komputer yang terhubung ke internet. Secara makna sebuah website adalah sekumpulan halaman informasi yang disediakan melalui jalur internet sehingga bisa diakses diseluruh dunia selama terkoneksi dengan jaringan internet tanpa terbatas ruang dan waktu. Web adalah suatu metode untuk menampilkan informasi di internet, baik teks, gambar, suara maupun video yang interaktif dan mempunyai kelebihan untuk menghubungkan (*link*) satu dokumen dengan dokumen lainnya (*hypertext*) yang dapat diakses melalui sebuah *browser*. Tahapan-tahapan pengembangan website dari awal

hingga akhir, adalah *planning, desing, developing / coding, upload web, promote, maintenance web*. Fungsi website secara umum adalah fungsi komunikasi, fungsi informasi, fungsi *intertainment*, dan fungsi transaksi.

2.2.7 MySQL

MySQL adalah salah satu aplikasi pengolahan database yang sering digunakan oleh banyak programmer, disamping sudah mendukung dalam pembuatan database yang berbasis *Client/Server*, juga dapat mengolah database dalam jumlah besar. SQL adalah bahasa standar untuk *query* yang digunakan untuk manipulasi data yang dibuat dalam berbagai DBMS (*Databas Management System*), salah satunya adalah MySQL. Beberapa keistimewaan yang dimiliki oleh MySQL adalah *Portability, Open Source, Multiuser, Peforment Tuning, Column Tuning, Command Dan Function, Security, Scalability Dan Limits, Connectivity, Localization, Interface, Cliens Dan Tools, Struktur Table*.